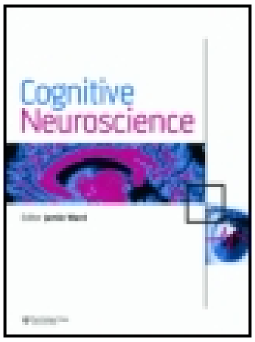

Cognitive Neuroscience

ISSN: 1758-8928 (Print) 1758-8936 (Online) Journal homepage: http://www.tandfonline.com/loi/pcns20

Behavioral measures of implicit theory of mind in adults with high functioning autism

Eliane Deschrijver, Lara Bardi, Jan R. Wiersema & Marcel Brass

COGNITIVE NEUROSCIENCE, 2015, http://dx.doi.org/10.1080/17588928.2015.1085375

Routledge
Taylor & Francis Group

Behavioral measures of implicit theory of mind in adults with high functioning autism

Eliane Deschrijver[1], Lara Bardi[1], Jan R. Wiersema[2], and Marcel Brass[1]

[1]Department of Experimental Psychology, Ghent University, Ghent, Belgium
[2]Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium

Theory of mind (ToM) research has shown that adults with high functioning autism (HFA) demonstrate typical performance on tasks that require explicit belief reasoning, despite clear social difficulties in everyday life situations. In the current study, we used implicit belief manipulations that are task-irrelevant and therefore less susceptible to strategies. In a ball-detection task, it was shown that neurotypical individuals detect a ball faster if an agent believed the ball was present. We predicted that adults with high functioning autism (HFA) would not show this effect. While we found a numerical difference in the hypothesized direction, we did not find a reliable group effect. Interestingly, the implicit ToM-index showed a strong negative correlation with both self-reported and observational measures of social difficulties in the HFA group. This suggests that the relationship between implicit ToM reasoning and the symptomatology of HFA might be subtler than assumed.

Keywords: Autism spectrum disorder; Implicit theory of mind; Explicit theory of mind; Belief reasoning; High functioning autism.

For many people, one of the most intriguing aspects of autism spectrum disorder (ASD) remains the social difficulties individuals with this disorder characteristically show. In everyday life social situations, for instance, individuals with ASD often demonstrate reduced eye contact, difficulties to infer other people's stand, and an egocentric focus within the normal reciprocity of social conversations (American Psychiatric Association, 2013). Some decades ago, Baron-Cohen and colleagues proposed that ineptitude at inferring mental states of others might lie at the basis of these difficulties (Baron-Cohen, Leslie, & Frith, 1985). They reasoned that a lack of Theory of Mind (ToM) would lead children with ASD to fail at imputing beliefs to others and at predicting their behavior (Baron-Cohen et al., 1985; Frith, 2001). To investigate this, they developed the Sally-Anne false belief task (Baron-Cohen et al., 1985; Wimmer & Perner, 1983), in which Sally places an object (typically, a ball) in a box and leaves the scene. After this, Anne notoriously relocates the ball and Sally returns. Participants with well-developed ToM capacities succeed in predicting Sally's behavior based on her false belief about the ball's location.

Since then, numerous studies have been devoted to scrutinize the belief attribution abilities of people with and without ASD. Neurotypical infants, for instance, have demonstrated basic indications of ToM as early as the age of two years or below (Baillargeon, Scott, & He, 2010; Onishi & Baillargeon, 2005; Senju, Southgate, Snape, Leonard, & Csibra, 2011; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007). Most children pass explicit tasks using manipulations of others' "false belief" at the age of four years. Children with ASD, however, are often outperformed, despite having at least an equivalent

Correspondence should be addressed to: Eliane Deschrijver, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium. E-mail: eliane.deschrijver@ugent.be

No potential conflict of interest was reported by the authors.

mental age (Baron-Cohen et al., 1985; Happé, 1995). Nevertheless, studies focusing on older children and adults with high functioning autism (HFA) often report apparently standard responses to stimuli that require belief processing (e.g., Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Peterson & Slaughter, 2007; Scheeren, de Rosnay, Koot, & Begeer, 2013). Since qualitative difficulties in social interaction persist for these individuals in everyday life, scientists assume that the use of compensatory strategies leads to ceiling effects in these tasks (Happé, 1995; Schneider, Slaughter, Bayliss, & Dux, 2015; Senju, Southgate, White, & Frith, 2009; Senju, 2013). As such, one could expect that these individuals would manage to conceal existing belief-processing problems only when they are aware of what is required to successfully perform the task.

Interestingly, researchers have recently developed a so-called implicit ToM task (Kovács, Téglás, & Endress, 2010), which contains stimuli with belief manipulations that are based on those of the Sally-Anne false belief task (Baron-Cohen et al., 1985), but merely require the visual detection of a ball's presence. Remarkably, the manipulations are considered to be implicit first because the agent and his beliefs are task-irrelevant (see also Clements & Perner, 1994; Low & Watts, 2013; Senju et al., 2011), and second because follow-up studies reported the participant's absence of awareness regarding the belief manipulations (Schneider, Bayliss, Becker, & Dux, 2012; Schneider, Lam, Bayliss, & Dux, 2012; Schneider, Nott, & Dux, 2014; Schneider, Slaughter, Becker, & Dux, 2014; Senju, 2013). The key finding in the implicit ToM tasks showed that representing the agent's belief influenced the ball-detection latencies of the participants (Kovács et al., 2010). Importantly, the paradigm showed implicit belief processing even within the looking times of infants as young as seven months. This adds to the assumption that the belief computations elicited by this task most likely happen spontaneous and without conscious deliberation (Kovács et al., 2010). In addition, recent neuroimaging studies showed that the task engages brain areas that are core to the explicit processing of other's beliefs (e.g., the right temporo-parietal junction and the medial prefrontal cortex; Kovács, Kühn, Gergely, Csibra, & Brass, 2014; Schneider, Slaughter, et al., 2014), suggesting that social cognitive processes most likely underlie implicit ToM.

Up until now, to our knowledge, only two preliminary studies have used implicit ToM tasks based on the Sally-Anne task to investigate ASD. These eye tracking studies found evidence of impaired implicit ToM in adults with HFA demonstrating absence of belief attribution to others in their looking times (Schneider, Slaughter, Bayliss, & Dux, 2013; Senju et al., 2009). However, it has never been investigated whether deficient implicit belief processing in ASD can be revealed by other measures than eyes movements. This is relevant, because social situations generally require not only to observe, but also to act upon the beliefs of others, and eye movement measures may be more strongly influenced by general attentional difficulties in ASD (Murray, 2010). Other studies that tried to measure ToM abilities in ASD populations in a less explicit fashion focused on stimuli that are far off from containing false belief manipulations and/or required deliberation on the stimuli's social content (Begeer, Bernstein, van Wijhe, Scheeren, & Koot, 2012; Callenmark, Kjellin, Rönnqvist, & Bölte, 2013; Rice & Redcay, 2014; Rosenblau, Kliemann, Heekeren, & Dziobek, 2013). It is likely that these paradigms required other cognitive processes or even explicit belief reasoning (see also Callenmark et al., 2013; Schneider et al., 2013). Therefore, in the current study, we adopted a version of the behavioral implicit ToM task (Kovács et al., 2010) in adults with HFA and a matched control group. We consider this task "implicit," because it investigates belief processes that are task-irrelevant and processed unconsciously (see also Schneider, Bayliss, et al., 2012; Schneider, Lam, et al., 2012; Schneider, Nott, et al., 2014; Schneider, Slaughter, et al., 2014; Senju, 2013). Much like in the Sally-Anne task, the scene shows an agent observing a ball either disappearing behind an occluder or leaving the scene. Accordingly, the agent respectively "believes" that the ball is present behind the occluder, or not. When the agent leaves the scene afterward, the participant's "belief" about the ball's presence is manipulated. Then, the agent returns and the occluder falls down, randomly showing the ball in half of the trials, while the ball is absent in the other half. Ball-detection latencies are measured as a function of both the implicit beliefs of the agent and of the participant.

As a main hypothesis, we expected that the ASD group would show deficiencies in implicit ToM belief processing. The original study (Kovács et al., 2010) showed that in conditions where the participant expected the ball to be absent, ball-detection latencies were speeded by the implicit representation of the agent's belief that the ball was present (as compared to a baseline condition in which both held the belief that the ball was absent). As such, implicit

ToM abilities can be investigated by evaluating the difference in ball-detection latencies between these two conditions (the "ToM-index"). Consequently, deficient implicit ToM skills in the ASD group would be expressed by a significantly reduced ToM-index, as compared with that of a matched control group. We also aimed to include measures of self-reported social symptom severity (the Autism Questionnaire (AQ) and Social Responsiveness Scale—Adult version (SRS); Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Costantino & Gruber, 2005; Noens, De la Marche, & Scholte, 2012) and of social dysfunctions shown in clinical observation (ADOS-interview; Lord et al., 1989) in the analyses. Because of its implicit character, we reasoned that the current ToM task would prove more sensible than explicit ToM tasks to detect belief processing difficulties of adults with HFA. We thus predicted that the diminishing of the ToM-index in the HFA group would reliably predict social difficulties of these individuals in daily life.

In a second and more exploratory approach, we wanted to investigate whether adults with HFA experience more difficulties than control participants when their own expectations about the ball's presence do not match the outcome. Previous research indicated that individuals with ASD find it more difficult than typically developing individuals to report their own *prior* false belief, suggesting that impairments in explicitly revising one's own belief exist in ASD (Russel & Hill, 2001; Williams & Happé, 2009). In addition, recent theoretical advances have associated ASD with general difficulties while processing mismatches between predicted and observed outcomes (Van de Cruys et al., 2014), suggesting that individuals with ASD have difficulties to disengage from reality, that is, from actual sensory input. We computed the "reality-bias" as a comparison of conditions in which the participant's belief about the ball's presence is false, against those conditions in which it is true. So, if individuals with ASD would have difficulties processing mismatches between a predicted and an eventual outcome, we would expect a larger reality-bias for the HFA group.

METHODS

Subjects

Twenty-three adults with HFA were recruited by means of our own research volunteer database and a recruiting announcement distributed by the Flemish Autism Association. Each participant was matched with a neurotypical control participant (CON) on demographics measures of age (±5 years), handedness, and gender. Participants in the CON group were screened for exclusion criteria prior to participation (neurological, psychiatric, sensory or motoric problems, and the use of psychiatric medication). All participants with HFA had received a formal diagnosis of ASD (including autistic disorder, Asperger's syndrome, and pervasive developmental disorder—not otherwise specified (PDD-NOS)) from an independent clinician or multidisciplinary team and reported no additional neurological disorders. Following ASD inclusion criteria for investigating HFA similar to other studies (e.g., Magnée, de Gelder, van Engeland, & Kemner, 2008; Zwickel, White, Coniston, Senju, & Frith, 2011), we included matched ASD-CON participant pairs in the data analyses for which the person with HFA attained a minimum ADOS score of 6 and scored maximum one point below cut-off on maximum one subscale of the ADOS-interview. We retained the data of 19 pairs of participants in our analyses (with 14 HFA participants meeting full ADOS criteria). We additionally excluded one control participant for having completed a professional training in perspective-taking. The remaining participant groups (HFA: n = 19, CON: n = 18) were well matched for gender (HFA: 13 males, CON: 12 males), handedness (right-handed HFA: 17 persons, right-handed CON: 17 persons), age (HFA: M = 32.95 years, SD = 6.26 years, range = 22–46 years; CON: M = 31.89, SD = 6.82, range = 21–46 years), and full-scale IQ score (HFA: M = 111, SD = 14.64; CON: M = 118, SD = 14.23). Chi-squared tests and t-tests confirmed that no significant demographic differences existed between groups. Due to missing data, the SRS questionnaire data of three individuals (HFA: 1; CON: 2) and the AQ questionnaire data of two participants with HFA could not be included. Individuals in the HFA group scored well above ADOS and autism cut-offs of the AQ on average (Lord, Rutter, DiLavore, & Risi, 1999; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). As one could expect, t-tests showed highly significant differences between the mean total dimensional scores on the SRS and on the AQ questionnaires. Participant characteristics and statistics are summarized in Table 1. All participants gave written informed consent and were financially compensated for their participation. The local ethics committee of the Faculty of Psychology and Educational Sciences approved the study.

TABLE 1
Participant details (***: Test is significant at the 0.001 level (2-tailed))

	HFA	*CON*	*t*	*p-value*
Number of male participants	13	12	N.A.	N.A.
Number of right-handed participants	17	17	N.A.	N.A.
Mean age (*SD*)	32.95 (6.26)	31.89 (6.82)	0.49	.63
Mean full-scale IQ (*SD*)	110.95 (14.64)	117.89 (14.23)	1.46	.15
Mean ADOS communication (*SD*)	2.58 (1.07)	N.A.	N.A.	N.A.
Mean ADOS social interaction (*SD*)	6.16 (2.17)	N.A.	N.A.	N.A.
Total score Autism Questionnaire (*SD*)	32.11 (8.44)	11.50 (4.05)	9.03	.00***
Total score Social Responsiveness Scale (*SD*)	159.33 (35.02)	93.69 (14.50)	7.28	.00***

Procedure

For both groups, the implicit ToM task was part of a larger battery of unrelated studies (not presented here) that were split up into two experimental sessions. This study was the first of the second session, which took place approximately three weeks after the first in a dimly-lit and sound-attenuated room. After another behavioral experiment, the session ended with the gathering of demographic data. If no standardized intelligence assessment had been performed within five years prior to participation, the participants' status as "high functioning" was derived from an IQ-score estimation using the KAUFMAN 2 short form Wechsler Adult Intelligence Scale III (full scale IQ ≥ 85; Wechsler, 1997; see Minshew, Turner, & Goldstein, 2005; for the use in adults with HFA). Participants with ASD completed the Autism Diagnostic Observation Scale-Module 4 (ADOS; Lord et al., 1989), administered by a formally trained researcher. Participants of both groups filled in self-report questionnaires measuring (social) autistic behavior: The AQ (Baron-Cohen et al., 2001) and the SRS (Costantino & Gruber, 2005; Noens et al., 2012).

Stimuli and task

We used an adapted version of the implicit ToM task (Kovács et al., 2010). Instead of a smurf as an agent (as in the original study), we employed the Buzz Lightyear figure of the *Toy Story* movies. We animated the movies using Autodesk's 3DS MAX software, and applied the original timing properties and storyline of the movies (see Figure 1). Replicating the original behavioral task of Kovács and colleagues (2010), we presented 10 trials for each of the four conditions in fully randomized order. Each of these conditions had two different outcomes (ball present/absent), resulting in eight different movies, each seen five times. As such, a total of 40 movies were presented with Presentation software.

In each movie, the beliefs of the agent (A) and of the participant (P) about a ball's presence behind an occluder were manipulated ("+" for a present ball, "-" for an absent ball). The first frame of each movie in the task shows a scene with a table and an occluder (see Figure 1). The movie starts with the agent entering the scene, and placing a ball on the table. The ball starts moving and hides behind the occluder. Then, four different scenarios could happen: In the P-A- condition, the ball leaves the scene while the agent is present. Then, the agent leaves, (implicitly) holding the belief that the ball is absent (A-). In the absence of the agent, the ball shows up in the scene once more, but then leaves the scene again. As such, the participant holds the belief that the ball is absent as well (P-). In the P-A+ condition, the ball is behind the occluder and the agent then leaves the scene. As such, the agent (implicitly) holds the belief that the ball is present (A+). After this, the ball emerges from behind the occluder, and leaves the scene. The participant now holds the belief that the ball is absent (P-), while the agent's implicit belief that the ball is present has become false. In the P+A- condition, the ball leaves the scene while the agent is present. Then, the agent leaves, (implicitly) holding the belief that the ball is absent (A-). In the absence of the agent, the ball shows up in the scene again, and then hides behind the occluder. As such, the participant holds the belief that the ball is present (P+), while the agent's implicit belief that the ball is absent has become false. Finally, in the P+A+ condition, the ball is behind the occluder while the agent leaves the scene. As such, the agent (implicitly) holds the belief that the ball is present (A +). After this, the ball emerges from behind the occluder, and then hides behind the occluder again. The participant thus also holds the belief that the ball is present (P+). After each of these four scenarios, the

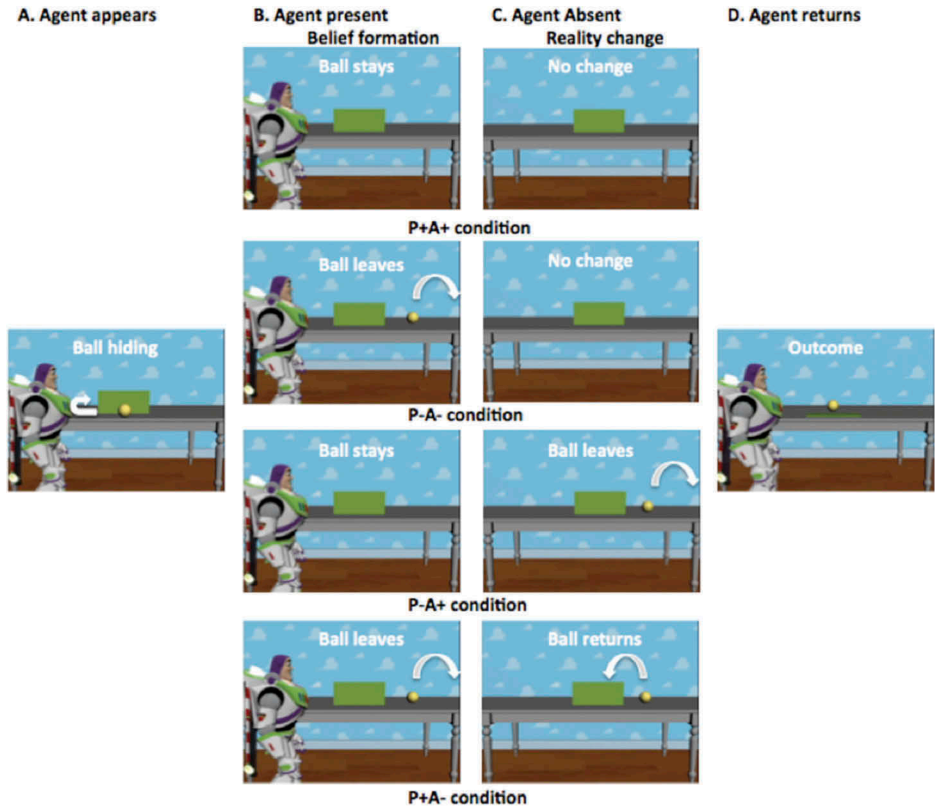

Figure 1. Design of the paradigm, based on the original article (Kovács et al., 2010). (A) The agent appears in all four conditions, and places the ball on the table. The ball then rolls behind the occluder. (B) In the A+ conditions (P+A+ and P-A+), the ball stays behind the occluder while the agent is present. In the A- conditions (P+A- and P-A-), the ball leaves while the agent is present. The agent then leaves the scene. (C) While the agent is absent, the ball reappears. In the P+ conditions (P+A+ and P+A-), the ball hides behind the occluder. In the P- conditions (P-A+ and P-A-), the ball leaves the scene. (D) The agent returns and the occluder falls down, randomly showing the ball in half of the cases (shown in the figure), while the ball is absent in the other half.

agent returns and the occluder falls down. We randomly manipulated the outcome of the movies (per condition): In half of the trials, the table now shows the ball. In the other half, the ball is absent.

Participants were instructed to perform a visual detection task, in which they had to press a button with their right hand as soon as they detected the ball when the occluder had fallen down. Ball-detection latencies were measured as a function of the beliefs of participant and agent. We only included the latencies in the conditions in which the participant actually had to press the key (present ball; five trials per condition). To preserve the participant's attention, they were required to press another button with their left hand as soon as the agent left the scene. Four practice trials were presented before the start of the experiment (discarded from analyses). Importantly, no instructions were given with respect to belief computations or the goal of the experiment. Previous research with these stimuli in neurotypical adults in our lab, and similar studies in other labs, showed that the reaction time pattern elicited by the task proved highly stable in a series of experiments (see also Kovács et al., 2010) and the stimuli have yielded very high unawareness rates regarding the belief manipulation (97% of all participants, as observed by Nijhof, Bardi, Wiersema, & Brass, in prep.; see also Schneider et al., 2013). For now, we have no theoretical reasons to assume that this should be different for individuals with HFA.

Based on our hypotheses and on previous research (Kovács et al., 2010) we focused on two dependent measures. First, we computed the ToM-index by subtracting the ball-detection latencies in the P-A+ condition (in which the agent believes the ball is present but the participant does not) from the baseline condition (P-A-) in which both the agent and the participant assume the ball is absent (Kovács et al., 2010). We consider the P-A- condition as a baseline condition, because research has shown that neurotypical participants react most slowly in this condition. The P-A+ condition, on the other hand, is most crucial for our ToM hypothesis: If the participant represents the agent's belief, ball-detection latencies should fasten in the P-A+ condition (compared to the baseline P-A- condition). The ToM-index should hence be positive (Kovács et al., 2010). If, however, the participant does not experience this facilitating influence, the ToM-index should be zero. Second, we computed a second index to which we refer to as the reality-bias. The reality-bias subtracts the conditions in which the participant believes that the ball is present (sum of P+ conditions), from the conditions in which participants believe the ball is absent (sum of P-conditions). While the reality-bias is not completely independent of the ToM-index, it emphasizes the role of the participant's own expectations about the ball's presence in relationship to the outcome, without taking into account the belief of the agent.

RESULTS

ToM task

First, we wanted to investigate the ToM abilities in our participant groups. The ToM-index of the CON group was positive, indicating that the P-A+ condition induced shorter reaction times than the P-A- condition (respective means $M = 461$ ms and $M = 499$ ms; $t(17) = 3.14$, $p < .01$). This reflects a facilitating influence of the agent's belief on ball-detection latencies of about 40 ms (Figure 2). From our hypothesis, we had anticipated that this difference would disappear for individuals with HFA. The ToM-index in the HFA group was, indeed, numerically almost zero ($M = 6$ ms) and a pairwise comparing the P-A- and P-A+ conditions in this group yielded a far-from-significant result ($t(18) = 0.27$, $p = .79$). However, when the ToM-indices of the groups were compared directly, they did not differ significantly ($t(28.3) = -1.31$, $p = .20$, degrees of freedom corrected). Despite the clear numerical group difference of more than 30 ms, we could not confirm the hypothesis that the HFA group would show a significantly reduced ToM-index, as compared to the CON group. Interestingly, further exploration of the data revealed that the HFA group showed more variability in the individual ToM-indices than the CON group (Levene's test = 5.35, $p < .05$), but only because of a greater variability in the crucial P-A+ condition. Indeed, a group difference in the variances of the response times was confirmed statistically for the P-A+ condition (Levene's test = 4.78, $p < .05$), while the variances in the other conditions did not show group differences (for P+A+: Levene's test = 0.00 $p = .98$; for P+A-: Levene's test = 0.02 $p = .90$; for P-A-: Levene's test = 0.53, $p = .47$). This suggests that the larger variability in the individual ToM-indices of the HFA group is mainly driven by reaction time differences in the crucial P-A+ condition, which was reported to reflect implicit belief processing (Kovács et al., 2010). Importantly, when testing whether the variation in the ToM-index in the HFA group was correlated with the severity of (social) autistic symptom, we observed reliable correlations with both self-report and clinical observation schedules: In the clinical group, the correlation between the ToM-index was highly significant for the AQ ($r = -.79$, $p < .000$), and marginally significant for both the SRS ($r = -.44$, $p = .07$) and the social subscale of the ADOS-interview scores ($r = -.44$, $p = .06$). Overall, this indicates that symptom severity, both in self-report (AQ and SRS) and in clinical observational measures

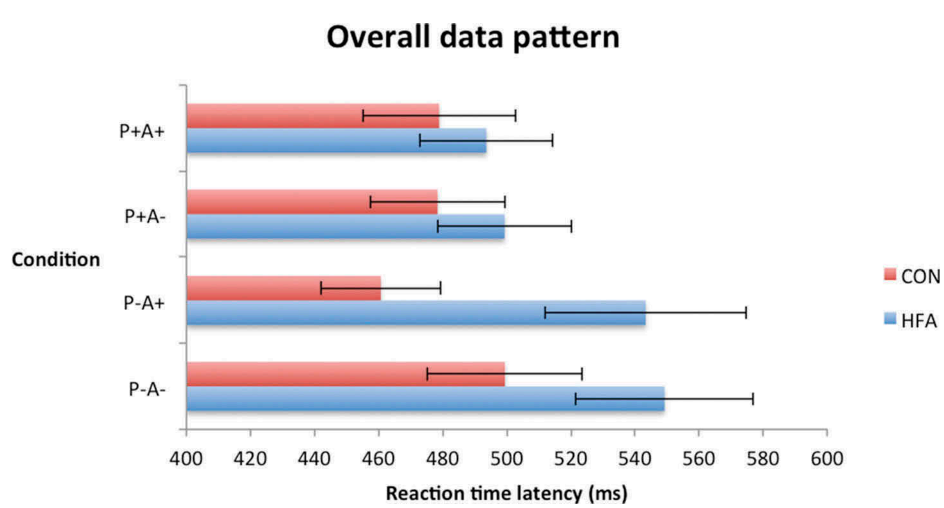

Figure 2. Data pattern per group and per condition. Standard error bars are noted.

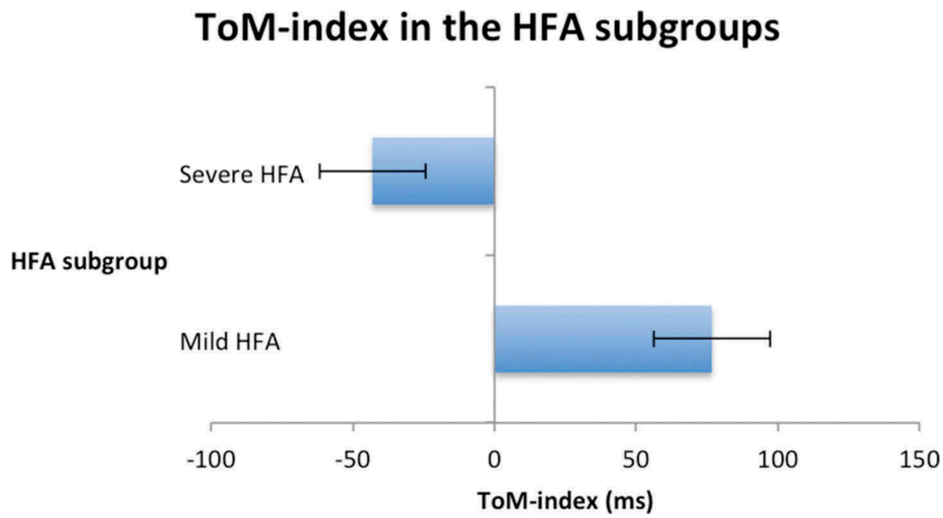

Figure 3. Plots of the mean ToM-index for high HFA and low HFA (n = 9 per participant subgroup). Standard error bars are noted.

(ADOS-interview), was negatively correlated with the ToM-index. This strongly suggests that the variability in the ToM-index of the HFA group was not due to noise (Figure 3). As such, we show that actual social difficulties of adults with HFA in daily life are reliably related to the ball-detection latencies of the ToM-index.

As an interesting post-hoc finding, inspection of the correlation plots showed that some individuals from the HFA group with more severe ASD symptoms reveal a *negative* ToM-index, while many with "mild" ASD symptoms showed a positive ToM-index, much like individuals in the control group (Figure 4). A negative ToM-index suggests that the ball-detection response in the P-A+ condition was *slowed down*, with respect to the response in the P-A- baseline condition. Our strong correlations suggest that the variation of the ToM-index in the negative direction seems to be relevant to HFA, especially since these correlations seem to be driven by the latencies in the P-A+ condition only. To explore the potential reversal of the ToM-index and its relation to autistic symptoms, we used the ADOS-interview total scores (the only demographic measure without missing values) to perform a median-split on the ToM data of the HFA group, resulting in a "mild HFA" and a "severe HFA" subgroup. An independent-samples t-test confirmed that the ToM-index differed between the subgroups ($t(16)$ = 4.30, p = .001; respective means 76 ms and -43 ms). Further comparisons showed significant differences between the crucial conditions for both the severe HFA group ($t(8)$ = -2.30, p = .05) and the mild HFA group ($t(8)$ = 4.42, p < .01). While these results are highly significant, they are only based on nine participants in each group and therefore should be treated with caution. With this in mind it is nevertheless interesting that the negative ToM-indices of individuals with severe ASD seem to counteract the positive ToM-indices of many individuals with mild

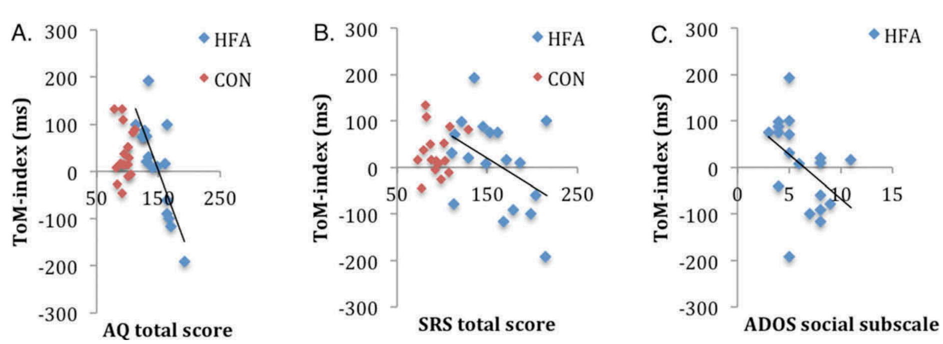

Figure 4. Correlational plots. A trendline is marked for the HFA data only. (A) The ToM-index and the AQ total dimensional score (including both groups). (B) The ToM-index and the SRS total dimensional score (including both groups). (C) The ToM-index and the ADOS social subscale score (for HFA only).

HFA, resulting in an averaged ToM-index of almost zero in the HFA group. Implications of this finding are discussed below.

Finally, we wanted to test whether participants with HFA experience more difficulties than control participants when their own expectations about the ball's presence do not match the outcome (as compared to when they do). Therefore, we compared the reality-bias for each group in an independent-samples *t*-test. The test showed a clear group difference ($t(25.58) = 2.99$, $p < .01$; degrees of freedom corrected, Levene's test: $F = 4.19$, $p < .05$), with a larger reality-bias in the HFA group than in the CON group (50 ms and 1 ms, respectively).

GENERAL DISCUSSION

Previous research demonstrated that adults with HFA usually perform at standard levels on explicit false belief tasks (e.g., Baron-Cohen et al., 1997; Happé, 1995; Peterson & Slaughter, 2007; Scheeren et al., 2013), though difficulties in social interactions persist in daily situations. It has, however, been argued that the discrepancy between social cognitive tasks and actual social abilities is consequent to compensatory strategies of adults with HFA that bring task performance to ceiling levels (Happé, 1995; Schneider et al., 2013, 2014; Senju, 2013; Senju et al., 2009; Williams & Happé, 2009). If this holds true, one would expect that adults with HFA show deficits in mentalizing tasks if the tasks do not allow for compensatory strategies (Clements & Perner, 1994; Low & Watts, 2013; Senju et al., 2011). In addition, one would expect the performance in these tasks to clearly relate to the social symptomatology of adults with HFA. The aim of the present study was therefore to test this hypothesis with the behavioral version of the implicit mentalizing task (Kovács et al., 2010). In this ball-detection task, a belief manipulation is implemented that is irrelevant to the task of participants. Accordingly, participants are usually unaware of the belief manipulation (Schneider, Bayliss, et al., 2012; Schneider, Lam, et al., 2012; Schneider, Nott, et al., 2014; Schneider, Slaughter, et al., 2014; Senju, 2013). We predicted that the HFA group should show impaired implicit processing of the belief of others, as reflected in a smaller ToM-index. We also expected a reliable relationship of this metric with actual social difficulties in self-report and observational measures. Furthermore, following recent theoretical advances on mismatching processing in ASD (Van de Cruys et al., 2014), we predicted that individuals with HFA would experience more difficulties than control participants in detecting a ball when their own expectations about the ball's presence do not match the outcome. Even though the group difference in the ToM-index was not reliable, our results revealed some interesting additional findings regarding implicit ToM processing in HFA.

Behavioral measures of implicit ToM in HFA

While we indeed found a numerically smaller ToM-index for the HFA group than for the CON group, this difference did not reach significance. Therefore, we did not find strong support for the hypothesis that the HFA group as a whole is impaired in implicit ToM processing. Indeed, many individuals with mild HFA showed ToM-indices within the range of those in the CON group. However, we observed that the ToM-index varied more strongly in the HFA group than in the CON group, due to variability in the P-A+ condition only, which is crucial for assessing implicit ToM abilities (P-A+ condition; Kovács et al., 2010). In addition, we found a negative correlation of the ToM-index in the HFA group with self-reported and clinical observation measures of social problems (respectively the AQ/SRS questionnaires and ADOS-interview). This suggests that in contrast to explicit ToM processing (Baron-Cohen et al., 1997; Happé, 1995; Peterson & Slaughter, 2007; Scheeren et al., 2013), social symptomatology is related to implicit belief processing in adults with HFA.

Furthermore, in an exploratory post-hoc analysis, we show that individuals in the HFA group with strong social symptomatology display *negative* ToM-indices, while those with mild social symptomatology show positive ToM-indices. Interestingly, the observation of a negative ToM-index seems to be primarily restricted to the HFA group. And, even though this analysis is based on a small number of participants in each group, it is highly reliable. Furthermore, it seems to be primarily driven by variation in the P-A+ condition, which is most crucially related to implicit belief processing. But what is the meaning of a negative ToM-index? It seems to indicate that a belief of the agent that is consistent with the outcome but incongruent to one's own expectation, leads to slowing of detection times in situations where the participant's own expectation about the absence of the ball is violated. In other words, while control participants profit from the belief of the agent when it is congruent with the outcome in a situation where their expectation is violated, this is not the case in the high HFA group. Instead, adults with severe HFA seem to be hindered by the fact that the agent held a belief that was true with respect to the outcome, but incongruent to one's own false belief.

This finding is difficult to reconcile with existing theories on ToM in HFA which would predict a reduction of the ToM-index in HFA but not a reversal. It should, however, be noted that the explicit ToM processing literature has been focusing almost solely on the processing of *false beliefs of others while holding a true expectation one's self*, while in the current task, the implicit *true belief* of the agent typically leads to the behavioral facilitation when confronted with a *violated expectation one's self*. As such, the implicit ToM task might yield insights that, from the stand of the explicit ToM literature, seem counterintuitive at first glance. While we think the reversed ToM-index is an interesting finding that of course needs to be replicated, we can for now only provide a tentative functional interpretation of this pattern. Interestingly, recent findings in the field of sensorimotor social cognition have suggested that individuals with ASD might experience difficulties while processing other-related information that stand in contrast to self-related information (Spengler, von Cramon, & Brass, 2009; Deschrijver, Wiersema, & Brass, 2015). These paradigms have shown strong theoretical and empirical links with ToM processing (Spengler, Bird, & Brass, 2010; Spengler et al., 2009) and with social skills in ASD populations (Deschrijver et al., 2015). From this perspective, the reversed ToM-index might not be due to the failure of representing the other agent's belief but rather due to confusion when the other agent has a belief that is inconsistent with the own belief. Interestingly, this confusion then only leads to a slowing of detection times if the participant had held a false belief one's self (P- conditions). This asymmetry between "positive" and "negative" false beliefs is consistent with a recent imaging study where it was found that brain activity during implicit belief processing was dependent on the participant's own expectations about the ball's presence (Kovács et al., 2014). If difficulties while processing mismatches exist in ASD (Van de Cruys et al., 2014), and ToM processing can be seen as a problem of solving these mismatches (Koster-Hale & Saxe, 2013), the double mismatch with respect to reality in the P-A+ condition might yield larger response times than the single mismatch in the P-A- condition, leading to reversed ToM-indices.

In sum, we could not provide support for the hypothesis that adults with HFA do not implicitly represent the agent. However, the ToM-index showed a linear relationship with actual social difficulties of these individuals, which has not been reported in explicit ToM tasks (Baron-Cohen et al., 1997; Happé, 1995; Peterson & Slaughter, 2007; Scheeren et al., 2013). Interestingly, we also observed *negative* ToM-indices for individuals with severe HFA, a finding presumably related to difficulties in processing self-other belief mismatches. In order to show such difficulties, however, the individuals with HFA first have to process the other person's belief. As such, the reversed ToM finding also supports the idea that individuals with HFA do implicitly represent the beliefs of the other agent. As such, our implicit ToM results are not in full accordance with eye tracking research that reported diminished implicit ToM processing in ASD at the group level. However, since results in these studies only rely on a small percentage of trials (e.g., around 10–20% of the data; Schneider et al., 2013), they might shed light on aspects other than the global implicit ToM abilities of these individuals.

Processing mismatches between own beliefs and the outcome in HFA

In line with our second hypothesis, we showed a clear difference in the way participants with HFA process instances where their own expectations about the ball's presence do not match the outcome. More specifically, we observed a larger difference between the P- conditions and the P+ conditions in the HFA group than in the CON group. This may indicate that the HFA group experienced a stronger-than-typical adherence to own beliefs/expectations, as has been reported in explicit ToM literature (Russel & Hill, 2001; Williams & Happé, 2009). Alternatively, it might suggest that these individuals experience disengagement problems from the unexpected outcome. Either way, the result adds to recent theoretical advances that have described ASD in terms of difficulties in processing mismatches between a predicted and an observed outcome (Van de Cruys et al., 2014).

Conclusion

In sum, our results suggest that the HFA group experienced difficulties in processing the mismatch between the ball's predicted and observed location, while the implicit processing of others' belief was not necessarily affected. Indeed, we did not find a group difference in the ToM-index metric. Rather, we show a more subtle relationship of ToM with the severity of the (social) autistic difficulties in both self-reported and clinically observed social measures of the HFA group. In our correlational analysis, we confirmed the hypothesis that the implicit ToM-index is related to actual social abilities of adults with HFA, a finding that has not been reported before in explicit ToM studies (Baron-Cohen et al., 1997; Happé, 1995; Peterson & Slaughter, 2007; Scheeren et al., 2013). In addition, we reported the post-hoc finding that individuals with severe HFA seem to represent the other agent, but this might sometimes hamper rather than facilitate the detection of the ball. Future research is certainly needed to replicate and further explore this finding. Given that implicit mental attribution is often regarded as a precursor of or even as proper ToM (Apperly & Butterfill, 2009; Kovács et al., 2010, 2014), the current results add significantly to the literature investigating social behavior in the autism spectrum. By no means, however, do we want to overgeneralize these findings in HFA to the larger population of individuals with ASD. Overall, the current findings stress the importance of including autistic trait measures and correlational results in the description of HFA/ASD data.

Original manuscript received 20 May 2015
Revised manuscript received 18 August 2015
First published online 13 November 2015

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970. doi:10.1037/a0016923

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118. doi:10.1016/j.tics.2009.12.006

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*(7), 813–822. doi:10.1111/jcpp.1997.38.issue-7

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46. doi:10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-spectrum Quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. doi:10.1023/A:1005653411471

Begeer, S., Bernstein, D. M., Van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism*, *16*(4), 357–366. doi:10.1177/1362361311434545

Callenmark, B., Kjellin, L., Rönnqvist, L., & Bölte, S. (2013). Explicit versus implicit social cognition testing in autism spectrum disorder. *Autism*, *18*(6), 684–693. doi:10.1177/1362361313492393

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395. doi:10.1016/0885-2014(94)90012-4

Costantino, J. N., & Gruber, C. P. (2005). *Social responsiveness scale (SRS). Manual*. Los Angeles, CA: Western Psychological Services.

Deschrijver, E., Wiersema, J. R., & Brass, M. (2015, May 22). *Action-based touch in high-functioning autism: Can compromised self-other distinction abilities link social and sensory difficulties in the autism spectrum?* 11th National Congress of the Belgian Society for Neuroscience, Mons.

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, *32*, 969–979. doi:10.1016/S0896-6273(01)00552-9

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66* (3), 843–855. doi:10.2307/1131954

Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, *79*, 806–848.

Kovács, A. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PLoS One*, *9*(9), 1–6.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*, 1830–1834. doi:10.1126/science.1190792

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule: Manual*. Los Angeles, CA: WPS.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Austism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*, 185–212. doi:10.1007/BF02211841

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*(3), 305–311. doi:10.1177/0956797612451469

Magnée, J. C. M., de Gelder, B., van Engeland, H., & Kemner, C. (2008). Audiovisual speech integration in pervasive developmental disorder: Evidence from event-related potentials. *Journal of Child Psychology and Psychiatry*, *49*(9), 995–1000. doi:10.1111/jcpp.2008.49.issue-9

Minshew, N. J., Turner, C. A., & Goldstein, G. (2005). The application of short forms of the Wechsler Intelligence scales in adults and children with high functioning autism. *Journal of Autism and Developmental Disorders*, *35*(1), 45–52. doi:10.1007/s10803-004-1030-x

Murray, M. J. (2010). Attention-Deficit/Hyperactivity Disorder in the context of autism spectrum disorders. *Current Psychiatry Reports*, *12*, 382–388. doi:10.1007/s11920-010-0145-3

Nijhof, A., Bardi, L., Wiersema, R., & Brass, M. (in prep.) Measuring mentalizing ability: a within-subject comparison between an explicit and implicit version of a ball detection task.

Noens, I., De la Marche, W., & Scholte, E. (2012). *SRS-A. Screeningslijst voor autismespectrumstoornissen*. Amsterdam: Hogrefe.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. doi:10.1126/science.1107621

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief and communication. *Child Development*, *60*, 688–700.

Peterson, C. C., & Slaughter, V. P. (2007). Social maturity and theory of mind in typically developing children and those on the autism spectrum. *Journal of Child Psychology and Psychiatry*, *48* (12), 1243–1250. doi:10.1111/j.1469-7610.2007.01810.x

Rice, K., & Redcay, E. (2014). Spontaneous mentalizing captures variability in the cortical thickness of social brain regions. *Social Cognitive and Affective Neuroscience*. Advance online publication. PMID: 24847726.

Rosenblau, G., Kliemann, D., Heekeren, H., & Dziobek, I. (2013). Approximating implicit and explicit mentalizing with two naturalistic video-based tasks in typical development and autism spectrum disorder. *Journal of Autism & Developmental Disorders*. Advance online publication. PMID: 25267068.

Russel, J., & Hill, E. L. (2001). Action-monitoring and intention reporting in children with autism. *Journal of Child Psychology and Psychiatry*, *42*(3), 317–328.

Scheeren, A. M., de Rosnay, M., Koot, H., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *The Journal of Child Psychology and Psychiatry*, *54*(6), 628–635. doi:10.1111/jcpp.12007

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of other's mental states. *Journal of Experimental Psychology: General*, *141*, 433–438. doi:10.1037/a0025458

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts implicit theory of mind processing. *Psychological Science*, *23*, 842–847. doi:10.1177/0956797612439070

Schneider, D., Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition*, *133*, 43–47. doi:10.1016/j.cognition.2014.05.016

Schneider, D., Slaughter, V. P., Bayliss, A. P., & Dux, P. E. (2013). A temporally sustained implicit theory of mind deficit in autism spectrum disorders. *Cognition*, *129*, 410–417. doi:10.1016/j.cognition.2013.08.004

Schneider, D., Slaughter, V. P., Bayliss, A. P., & Dux, P. E. (2015). What do we know about implicit false-belief tracking. *Psychonomic Bulletin & Review*, *22*(1), 1–12. doi:10.3758/s13423-014-0644-z

Schneider, D., Slaughter, V. P., Becker, S. I., & Dux, P. E. (2014). Implicit false-belief processing in the human brain. *NeuroImage*, *101*, 268–275. doi:10.1016/j.neuroimage.2014.07.014

Senju, A. (2013). Atypical development of spontaneous social cognition in autism spectrum disorders. *Brain & Development*, *35*, 96–101.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, *22*(7), 878–880. doi:10.1177/0956797611411584

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science*, *325*, 883–885. doi:10.1126/science.1176170

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*, 587–592. doi:10.1111/j.1467-9280.2007.01944.x

Spengler, S., Bird, G., & Brass, M. (2010). Hyperimitation of actions is related to reduced understanding of other's minds in autism spectrum conditions. *Biological Psychiatry*, *68*(12), 1148–1155. doi:10.1016/j.biopsych.2010.09.017

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Human Brain Mapping*, *30*, 3704–3718. doi:10.1002/hbm.v30:11

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*(7), 580–586. doi:10.1111/j.1467-9280.2007.01943.x

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649–675.

Wechsler, D. (1997). *Wechsler adult intelligence scale-III (WAIS-III)*. *Nederlandstalige bewerking*. Amsterdam: Harcourt Test.

Williams, D. M., & Happé, F. (2009). What did I say? Versus what did I think? Attributing false beliefs to self amongst children with and without autism. *Journal of Autism and Developmental Disorders*, *39*, 865–873. doi:10.1007/s10803-009-0695-6

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. doi:10.1016/0010-0277(83)90004-5

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*(3), 331–335.

Zwickel, J., White, S. J., Coniston, D., Senju, A., & Frith, U. (2011). Exploring the building blocks of social cognition: Spontaneous agency perception and visual perspective taking in autism. *Social Cognitive and Affective Neuroscience*, *6*, 564–571. doi:10.1093/scan/nsq088